**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| **Name: Avinash Kumar Patni**  **Email:** [**avinashpatni0000@gmail.com**](mailto:avinashpatni0000@gmail.com)  **Contribution:**   1. Importing libraries 2. Import of csv datasets to google colab 3. Drive mounting  * Data cleaning - * Check the Null value * Null value replacing * Merge of two datasets * Handling duplicate value and multiplying data set * Data Exploration--Univariate & Bivariate Analysis * Exploring User review data frame * Handling the error and NaN values in the User reviews * Data Visualization on play store data * Data Visualization on User Reviews * How Content Rating affect over the App  1. Conclusion 2. Summary 3. Challenges and future work |
| **Please paste the GitHub Repo link.** |
| GitHub Link:- https://github.com/avinashpatni000/App\_Store\_EDA\_Project |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**  Using the dataset, developers can get insights into customer demands and make their product more popular. The dataset can also be used to determine whether the app's original ratings match its predicted rating in order to determine whether it is performing better or worse than other apps on the Play Store.  For app developers, Data Science is a very useful tool for building applications that focus on particular categories discussed in this analysis.  I merged two datasets, play store app review and user data review, in which play store app review has more columns than user review.  The app from play store is mainly classified by categorical type of data in which we do the EDA on each app by categorical value.  The approach in this I have to group-by with features of app that give me the major aspects. The most critical in this project is deal with the mixture of data it means it consists of string types of data The approach behind this make data type of numerical feature into float. This will assist me in evaluation of this features.  The conclusion of the entire project is the careful handling of data. A single mistake or ignorance of the data will lead to more problems, it will happen as if you are stuck in maize. So be careful with each step.  I type some question try to give the solution of that question that was the key purpose for whole my project. |
|  |